Dynamic exposure control in color scanners

Giordano Beretta, Computer Peripherals Laboratory

HPL–97–30
January, 1997

image processing, color facsimile, scanning, sheet-fed, shading correction, read data normalization, exposure enhancement

The lightness values of white papers cover an approximate range of fifteen jnd. The tone range of scanners is adjusted for the lightest possible substrate. Therefore, a scan is usually preceded by a preview operation in which the image is subsampled and the data is analyzed to determine the actual tone range. In color facsimile and sheet-fed scanners that do not buffer the entire image, such an operation is not possible.

We present a technique in which statistical methods are used to estimate the tone level of the paper. This estimate is used to set the parameters for a tone reproduction curve. The technique is incremental, the statistical data is gathered during the scan. While the scan progresses, the estimate is refined based on the increased amount of data avail able from the accumulated histogram. This has also the advantage that artifacts due to lamp warming during slow scans (typical for color facsimile) are automatically compensated.

1 Introduction

1.1 Problem statement

Although the light source in a scanner is more or less constant, papers reflect light to a varying degree. Scanners are designed to deliver a monotonically increasing signal for an increasingly bright paper. The response range is designed to encompass all possible papers. Typically, the response of a scanner sensor is linear with intensity.

To achieve a large dynamic range, the human sensorium compresses the response to stimuli. *Weber's law* (1834) states that the change in stimulus intensity that can just be discriminated is a constant fraction of the starting intensity of the stimulus. By integrating Weber's law over a series of stimulus intensities, we obtain *Fechner's law* (1860): the sensation magnitude is proportional to the logarithm of the intensity of the stimulus in units above absolute threshold. In particular, Fechner's law applies also to the *human visual system* (HVS). Due to this logarithmic compression and the human tendency of exploiting the sensorium's bandwidth, users of scanners an other digital input devices like those used in digital photography and video, will capture originals spanning a large dynamic range.

The designer of the digital device is confronted with the problem that the linear sensor will cause a poor response in the dark areas. To minimize artifacts, it is important to utilize carefully the sensor's dynamic range, i.e., not to waste any portion of the signal bandwidth. We report on a method to utilize efficiently the dynamic range in a sheet-fed scanner, for applications like color facsimile transmission and networked scanning stations.

1.1.1 Fechner's law repealed

Fechner's law is important because a common solution to the problem is to apply a logarithmic transformation to the signal. The typical technology used for sensors is a *charge coupled device* (CCD), which has linear response, hence the transformation is applied after the *analog-to-digital* (*AD*) conversion and proper utilization of the available range is still important.

More recently, in the 1950s, Fechner's law has been repealed in favor of *Stevens' power law*: percepts are proportional to the n-th power of stimulus intensity. The perceptual response can be linear ($n = 1$, e.g., length), compressive ($n < 1$, e.g., brightness), or expansive ($n > 1$, e.g., pain). In our application we consider brightness; scaling experiments suggest the use of $n = {}^{1}/_{3}$. More recent studies in the late 1950s and early 160s by Jameson and Hurvich indicate that Stevens' power law is not accurate, especially for complex fields. Bartleson and Breneman [2] in 1967 presented data indicating that for photographic images brightness does not correlate well with a simple power function of luminance. Their model includes both a power term and an exponential-decay term. Hunt in his appearance model uses a hyperbolic response. For current systems the departure from Stevens' law is not numerically significant and we need not consider them for exposure control, but they are important for the design of tone reproduction curves.

Percepts like the brightness are not useful for building scanners because they are not physical quantities and hence cannot be measured with an instrument. Because of the confusing and inconsistent terminology, especially in the patent literature, we set forth by developing the terminology, based on international standards.

1.2 Terminology

1.2.1 Predicting lightness

The attribute of a visual sensation according to which an area appears to exhibit more or less light is called *brightness*. Variations in brightness range from "bright" to "dim." *Lightness* refers to the brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting. Variations in lightness range from "light" to "dark."

As Bartleson and Breneman [2] note, lightness should not be confused with relative brightness. *Absolute brightness* is a response-oriented concept dealing with the absolute magnitude of intensive perception. *Relative brightness* also deals with the magnitudes of intensive perceptions but relative to the magnitude of a perception elicited by some tangible reference. Lightness is a stimulus-oriented concept which deals with the reflectance of stimulus objects; i.e., the perceived capacity of an object to direct light to the eyes. Lightness involves apperception as well as perception.

An example of a bright area is the *perfect reflecting diffuser*, an ideal isotropic diffuser with a reflectance equal to 1. A *diffuser* is a device used to alter the spatial distribution of radiation and depending essentially on the phenomenon of diffusion (scattering). *Isotropic diffuse reflection* refers to diffuse reflection in which the spatial distribution of the reflected radiation is such that the radiance or luminance is the same in all directions in the hemisphere into which the radiation is reflected.

Percepts like brightness and lightness are attributes of visual sensations. Of practical value are *correlates*, physical quantities that correlate well with the percepts.

Psychophysics is the discipline that deals with establishing correlates. For the problem at hand, the interesting psychophysical data is the photopic spectral luminous efficiency function $V(\lambda)$. For a monochromatic radiation of wavelength λ, the *spectral luminous efficiency* is defined as the ratio of the radiant flux at wavelength λ_m to that at wavelength λ such that both radiations produce equally intense luminous sensations under specified photometric conditions and λ_m is chosen so that the maximum value of this ratio is equal to 1. The sensations are those for a standard observer.

From a mathematical point of view, $V(\lambda)$ is a weight or measure. If the spectral power distribution of a stimulus is integrated (or summed) with this measure, the *luminance* is obtained, up to a constant factor. The luminance allows us to predict whether for a standard observer two stimuli of arbitrary spectral composition will appear to have the same intensity under the same viewing conditions. This prediction can be made evaluating a physical measurement.

The *Y tristimulus value* is a colorimetric quantity proportional to the luminance and can be read directly with the aid of almost all color measurement instruments. The Y tristimulus value can be absolute or relative, the latter being more common. When Y is evaluated on an absolute basis in candelas per square meters, it represents the luminance of a color and provides a basis for a correlation with the perceptual attribute of brightness. When Y is evaluated such that, for the similarly illuminated and viewed perfect reflecting diffuser, $Y = 100$, then Y is equal to the reflectance factor expressed as a percentage. This provides the basis for a correlate of the perceptual attribute of lightness. When the reference white is not the perfect diffuser, the ratio Y/Y_n is of interest, where Y_n is the value of Y for the *reference white*.

To find a good correlate of lightness, we bring into play Stevens' power law. By "good" we mean a measure that is approximately perceptually uniform. Because a power law produces negative values, for which there would be no perceptual basis, a correction is introduced by approximating that range by a linear function through the origin. The resulting quantity is called the *CIE 1976 lightness L** and is defined by

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16 \quad \text{for} \quad Y/Y_n > 0.008856 \tag{1}$$

and

$$L^* = 903.3\frac{Y}{Y_n} \quad \text{for} \quad Y/Y_n \leq 0.008856 \tag{2}$$

1.2.2 Color

A *trichromatic system* is a system for specifying color stimuli in terms of tristimulus values, based on matching colors by additive mixture of three suitably chosen reference stimuli. A set of three color stimuli on which a trichromatic system is based is called *reference color stimuli*. In the CIE (International Commission on Illumination) standard colorimetric systems, the reference color stimuli are represented by the symbols [X], [Y], [Z].

The *tristimulus values* of a color stimulus refer to the amounts of the three reference color stimuli, in a given trichromatic system required to match the color of the stimulus considered. In the CIE standard colorimetric systems, the tristimulus values are represented by the symbols *X, Y, Z*.

2 Scanner pixel flow

2.1 The scanning process

A *scanner* is a device that maps an *original* document to a *sample map* in a computer memory. The name "scanner" is because the device scans the original in a zig-zag pattern or line by line from top to bottom. In each atomic scanning operation, the intensity at a *photosite* on the original is mapped into a *pixel* or *picture element* (pel) of the sample map. This map is a discretization, in which real values are thresholded or binned into natural values. This results in quantization errors, which are perceived as *contouring* when the effect is undesired or *posterization* when the effect is intentionally used for artistic reasons.

The embodiment of the mapping function is called *channel*. In a monochromatic scanner—usually called *grayscale scanner*—there is one channel, while in a *color scanner* there are usually three channels allowing the representation of each photosite in a trichromatic system. The number of bits allocated to each channel depends on the amount of contouring that is acceptable in a given application. In the special case where only one bit of information is used to represent a photosite, the scanner is called a *black and white scanner* and the sample map is called *bitmap.*

2.2 Color and grayscale scanners

In modern scanners it is customary to represent the sample in a *color space* related linearly to a CIE space. Instead of "color space," the term *color model operator* is sometimes used, in reference to the colorimetric transformation applied to the data. A popular space is NTSC/RGB, which over time will probably be replaced by sRGB, which relates more closely to actual physical devices.

A color scanner can easily be used to produce grayscale images. To this purpose, the NTSC/RGB signal is transformed to CIE/XYZ and only the Y tristimulus value is considered. For illuminant C, the formula is

$$Y = 0.299R + 0.587G + 0.114B\,. \tag{3}$$

Because more than half of the contribution comes from the green channel, sometimes the grayscale image is obtained by simply discarding the red and blue channels.

To avoid contouring, the data-path width in pixels is often larger before the color space transformation than after. Because of Fechner's law, the color transformation is often combined with a power transformation of the form

$$y = x^{\gamma} \tag{4}$$

with $\gamma = 2.2$ being a common value. This allocates more bits to the dark portions of the image for a better rendition in the shadows. When Eqn. 4 is applied, we say that the conversion uses *gamma-corrected* R, G, B signals rather than tristimulus R, G, B values. The next step in sophistication is to consider Steven's law and output the scanner's signal in a space like CIELAB. This is done in some high-end desktop scanners.

2.3 Low-level scanner adjustments

Although the CCD sensors have a good linearity, they have two problems. The first problem is that some current leaks through the sensor even when no light impinges upon it. This is called the *dark current*. The second problem is that in many desktop scanners, instead of scanning the original's photosite with a single detector called *flying spot*, a number of sensors is manufactured on a single chip. The advantage is that the resulting *CCD array* can scan an entire line without movement. The disadvantage is the different CCD elements have a slightly different response to light, called the *photo response non-uniformity* or in short *PRNU*.

The dark current problem is solved by applying an offset to the signal, and the PRNU problem is solved by adjusting the gain factor individually for each CCD element. The parameters for this adjustment can easily be computed after the *analog-to-digital converter* (ADC). A white reference surface is positioned at the beginning of the scanner platen (scan line –1) and scanned first with the lamp still off to determine the offset and then with the lamp on to determine the gain for the individual CCD elements.

The signal compensation for these two effects can occur either in the digital domain, where the parameters are determined, or in the analog domain by adjusting the operation parameters of the ADC. Concomitantly the PRNU correction also compensates for possible non-uniformities in the lamp.

Unfortunately no uniform terminology is used in the literature for the dark current and PRNU adjustment. The following terms have been found: normalization, shading correction, PRNU- and dark-voltage compensation, gray level correction. The term shading correction is used also to designate the compensation for non-uniformity in the light source and in the optics[6] as well as to designate all gray level adjustments.

2.4 Summary

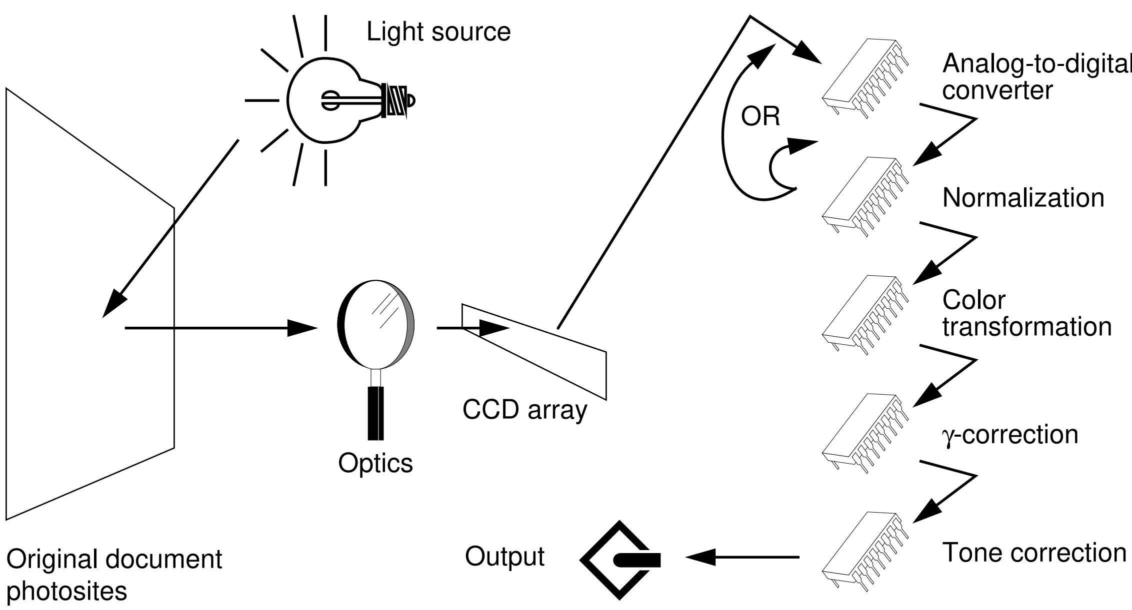

Figure 1. Conceptional pixel flow in a typical scanner. The order of the processing blocks on the right can be different.

The scanner pixel flow is summarized in Fig. 1. Light from a source impinges upon an original document. Each photosite is imaged onto a CCD element through an optical system. The scanner electronics comprises an analog to digital converter, optionally with adjustable gain and offset. The digital signal is subsequently “matrixed” in a color transform stage and the tone levels are re-mapped before the signal is output to a host computer. In the case of a color facsimile machine, the signal is output to stages for color encoding and data compression before being passed to the datacom logic.

3 Exposure control

Conceptionally, the adjustment for the background comes after the low-level adjustments described earlier. This adjustment can take place in the host computer or preferably in the scanner, where it is usually combined with the reduction of the pixel depth from the internal scanner value (usually between 9 and 12) to the external 8 bit value. The adjustment takes the form of a tone correction curve for each channel. Although this has nothing to do with the exposure duration of a scan line, we call it *exposure control*, because it makes the image lighter or darker. A similar term used in the patent literature is *exposure enhancement*.

It is possible to combine the exposure control with the electronics for the normalization. If this is done by altering dynamically the offset and gain in the ADC, it is probably sufficient to implement the scanner in just 8 bits per channel instead of 10. Alternatively, since the exposure control is the same for all CCD array elements, it could also be controlled by changing the power of the light source (possible problem: this might change the correlated color temperature), or by adding a mechanical iris or an electronic shutter to the optical system.

3.1 Paper lightness

As seen earlier, L^* is a colorimetric quantity that correlates well with the perceived lightness of a paper. This measure is scaled so that the perfect reflecting diffuser has value 100; one unit corresponds approximately to a *just noticeable difference* (jnd). Table 1 below shows the L^* and Y tristimulus values for some typical papers, measured for CIE illuminant D_{65}:

Medium	L*	Y
perfect reflecting diffuser	100	100
HP Premium paper	97	94
Cromalin proof	96	89
recycled copier paper	95	88
photographic AgX paper	93	82
3 years old acid paper	90	75

Table 1. Lightness and stimulus luminance of some typical papers compared to the perfect diffuser.

Medium	L*	Y
European recycled paper	87	71
fresh newspaper (outside print area)	85	66
4 months old newspaper	83	63
fresh newspaper (inside print area)	81	59

Table 1. (Continued) Lightness and stimulus luminance of some typical papers compared to the perfect diffuser.

Table 1 indicates that the perceived lightness difference of typical papers spans a range of 16 jnd units. In a linear intensity space where sensors operate, such as the *Y* tristimulus value, the range is from 94 for the HP 51634Y Premium Inkjet Paper to 59 for the newspaper in the print area.[*] We call the range of values present in the original document the *tone gamut*. Typically the lowest value in the gamut corresponds to black ink and the highest value corresponds to the paper background. We call the range of values for which a unique signal is generated by the lamp and sensor system in a scanner its *dynamic range*.

3.2 Impact on scanner design

A system is usually designed so that its dynamic range contains an original with the darkest ink on the lightest paper. Because of fluorescent substances often used in paper, the lightest value can be larger than the value of the perfect reflecting diffuser. We call this dynamic range the *worst case*.

When the image is output on a display monitor or a printer, this device will also be calibrated for the worst case, i.e., the lightest possible value will be used to represent the worst case lightest pixels. In a system that reproduces lightnesses relative to the perfect reflecting diffuser, an image scanned from a darker background such as a photograph will be reproduced by using ink to make the background darker (or the signal to a display monitor will be reduced) because the value without ink (or full display monitor signal) will be reserved to the worst case.

In a grayscale scanner based on the shortcut of using only the green channel to produce the image, according to Eqn. 3, the proportion of the New York Times background's contribution to the sample map is 0.587 * 0.59 = 0.346 versus 1 for the

*. Fresh newspaper means a just delivered copy of the New York Times measured immediately after unpacking. We measured the on bottom (i.e., innermost) margin of the first page using the reminder of the newspaper as backing, making sure no area has ink. The old newspaper was exposed to fluorescent indoor light (three-band, F11) for 4 months.

perfect diffuser in color. In other words, only a third of the signal bandwidth is utilized.

The HVS adjusts for the background or surround. This is why lightness—brightness relative to an object area perceived to be white—is a more important perceptual attribute than brightness in the case of grayscale and color reproduction. If a reproduction system does not map the input paper lightness into the output paper lightness, the appearance of the facsimile will be judged as inferior (e.g., smudgy or grayish) by a human observer.

3.3 Proposed solution

Statistical methods are used to estimate the reflectance of the paper and align the tone gamut to the scanner's dynamic range during a single scan. The estimate is refined during the scan. This refinement increases the robustness with respect of unfavorable originals (e.g., with a dark border) and at the same time can also take into account such phenomena as the change of the sensor's response due to thermal effects during slow scans. Slow scans are typical in color facsimile applications or in network scanning over slow networks.

4 Prior solutions and their disadvantages

The standard solution adopted in many desktop scanner applications and digital copiers is to perform a first preview scan at a low resolution. The so obtained image is analyzed to determine the original's paper color (or background value) and other parameters. The values are then used to set the scanner controls, after which the final scan is performed. The controls consist typically of the tone correction curve's parameters as shown at the right side in Fig. 1.

In applications such as sheet-fed scanners and color facsimile, it is not possible to perform a preview scan. In particular, in the case of a color facsimile machine, the image is transmitted during the scan and concomitantly printed on a remote machine, precluding any form of post-processing.

As mentioned earlier, a solution sometimes used is to assume a worst case or a typical case. If the assumption errs towards to high a lightness, the scanned image will not have a white background, yielding a smudgy appearance, decreasing the performance of image compression algorithms, and consuming more ink or toner when the image is printed. If the assumption errs towards to low a lightness, detail will be lost in the image's highlights, yielding a washed-out appearance.

5 Description of the method

In the present method no spatial information on the image is collected. Instead a histogram of the occurrence of each tone level is accumulated during the scan. The background color of the image, i.e., the tone value of the paper is estimated by analyzing statistically this histogram. When this value is known, the parameters for a tone reproduction curve can be calculated.

The shape of the TRC is not altered by these parameters, i.e., the exposure can be controlled independently of the tone reproduction. This process is repeated from time to time during the scan to assure robustness and to compensate for drifts in the sensor's sensitivity or in the lamp performance.

5.1 Tone levels in a document

Fig. 2 shows the red channel histogram of a color photograph. More precisely, a photographic color print is scanned and for each level in the 8-bit red channel the number of pixels (frequency) with that tone level is counted. A low tone level value indicates a dark tone and a high value indicates a light tone.

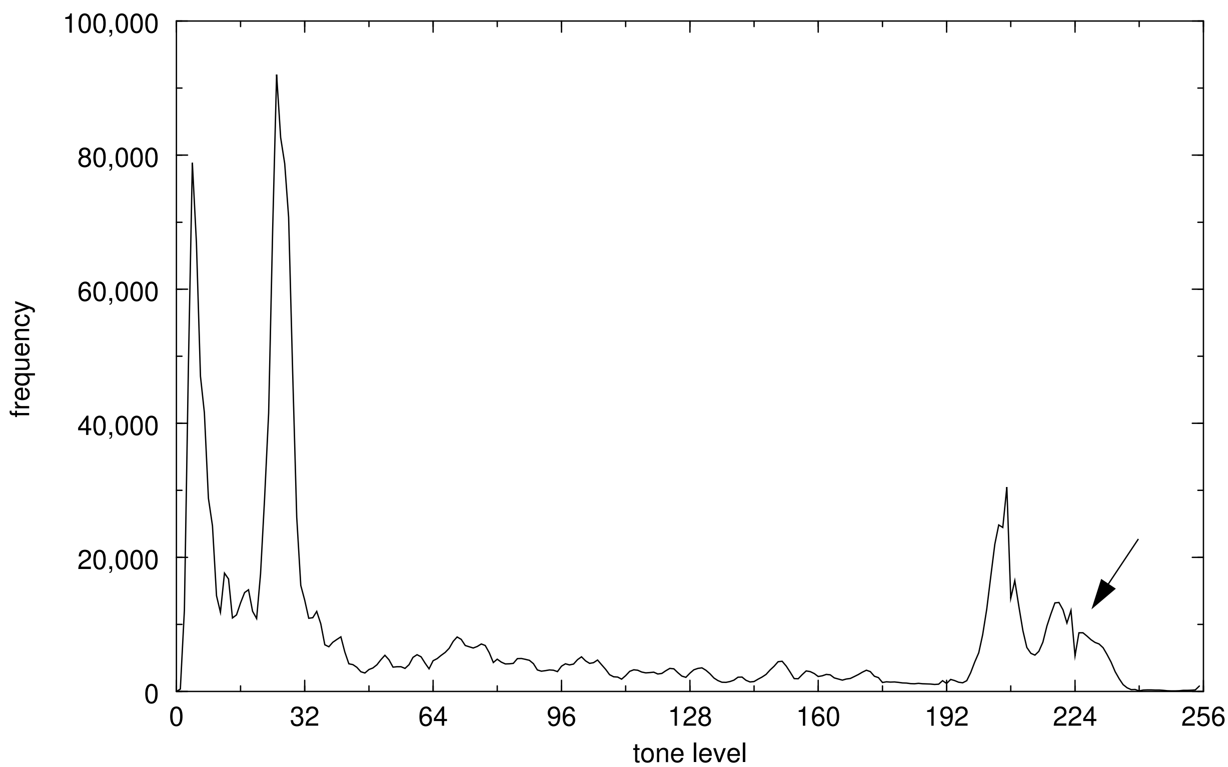

Figure 2. Histogram of a typical color photograph on silver halide paper.

In the dark region of the histogram, the first peak corresponds to black pixels such as in text and the second peak corresponds to shadows. In the light region, the left peak corresponds to the highlights and the right peak corresponds to the paper background.

Our aim is to estimate the tone level at which the peak for the paper occurs. The exposure control consists in building a tone reproduction curve that maps this peak tone level into the value 255 representing white. This procedure aligns the image's tone gamut with the scanner's dynamic range, so that a predictable and distinct signal is generated for each tone level.

5.1.1 Binning errors

The arrow in Fig. 2 indicates a common problem that occurs in the analog to digital conversion, namely a *binning error*. The device converts the tone level of the photosites with this particular same stimulus luminance with a higher probability into the value 223 than in the value 224. Because of this non-monotonicity, the peak cannot be found with local methods, i.e., by scanning the histogram from right to left until a maximum is reached. The binning errors can be so severe that smoothing the data with various methods such as moving window averages (MWA) or applying local least-squares fits (Savitzky-Golay smoothing filters) does not allow for robust estimations.

5.2 Tone levels of paper

We have examined statistically the light portions of a number of images by computing their central moments. Fig. 3 below shows the histogram for a sheet of blank HP Premium Ink Jet paper. This paper is typical in that it contains fluorescent substances that interact with peaks from the fluorescent light source used in the HP ScanJet IIc scanner. This produces a greater response in the green channel and a reduced response in the red channel. For this reason we can limit our method to examining the red channel, because a tone map that maps the red's peak into 255 will also map the blue and green peaks to that value, due to minmax clamping.

Notice that if we select carefully the lower boundary for the light portion of the histogram, the curves are relatively symmetric. The *mean* estimates the value around which central clustering occurs. Strictly speaking, this is true only for normal distributions, but due to the binning error problem, the mean is a more robust estimator than the mode or the median. We make an allowance for the non-normal distribution by using the maximum likelihood estimator.

As Fig. 3 suggests, we have to estimate not only location of the peak but also the variability of the data around the expected value to obtain a full shading correction. Our experiments have shown that the third (skewness) and fourth (kurtosis) central moment vary widely from sample to sample. Therefore the variance (or the standard deviation) is not a good estimator for this variability. Fortunately other estimators are available. The *average deviation* or *mean absolute deviation* provides for a more robust estimation of the variability in the data. As before, also in this case we use the maximum likelihood estimator to allow for the non-normality of the distribution.

5.3 Estimating the exposure

5.3.1 Parameter choices

We determine the exposure correction for a scan by accumulating histogram data on the tone levels in the red channel and then estimating the paper's tone level by adding d average deviations to the mean value (to avoid background noise). For the samples

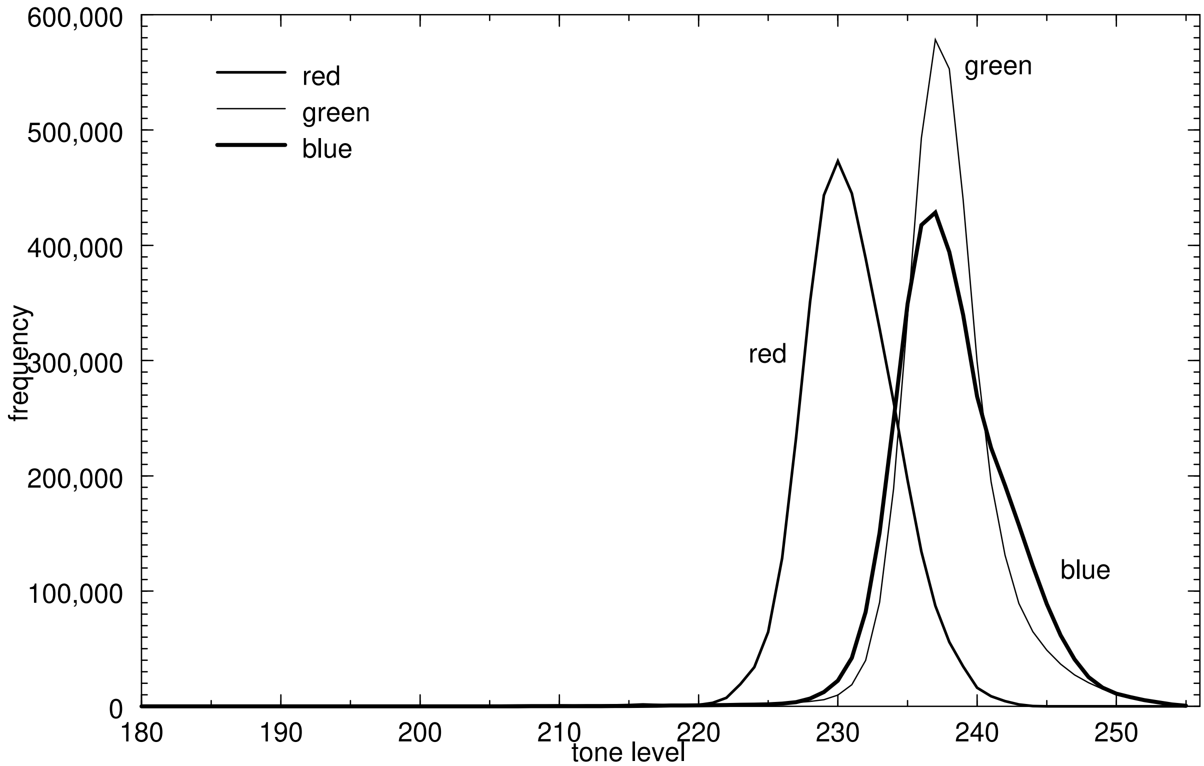

Figure 3. Light end of the histogram for a blank sheet of HP Premium Ink Jet paper.

we examined $d = 2$ was a good choice, but analysis of a larger number of different images might suggest a different value. Estimating the paper's tone level means that our method is independent of the image and the particular artistic use of the tone range used to express it. Two issues still remain to be discussed: when the analysis occurs and how the tone map is created.

5.3.2 Dynamically updating the estimate

Updating the histogram for each scan line is computationally inexpensive and is a feasible solution. The background color may change rapidly at the beginning due to paper skew and possible margins or rules in the original artwork. However, during most of the scan the paper's tone level will not change much and it is æsthetically undesirable to change the tone reproduction curve too harshly. From a statistical point of view, it is superfluous to examine every pixel to determine the paper's tone level because an image does not consist of uniformly distributed random pixels.

For many cases it might be possible to use a simple scheme, such as "examine each pixel in the first 20 scan lines and then examine the histogram every 100 lines." Such a simple method is not robust. Assume for instance the scanned original is a ruled paper with a rule spacing corresponding to 50 scan lines. If the paper is aligned unfavorably, it may well be that the samples are taken exactly at the photosites corresponding to the rules, resulting in an incorrect exposure.

In our investigations we observed thermal effects that have also been reported in the patent literature.[9] Due to these effects, the average deviation increases as we scan down the page, especially in an application such as color facsimile where the scanning speed is bound by the communication line speed. Consequently, the histograms are not identically distributed across scan lines. At first this might seem a problem, however, consider the following facts:

1. The maximum is "pushed out" of the dynamic range to avoid background noise (we mentioned earlier that we add d average deviations to the mean).
2. We look only at the histogram's light portion to avoid failure in the case of originals with a border or other artistic uses of the tone range in an image or document.
3. We assume that a picture or graphic will never bleed out of the image edge. This is a reasonable assumption because the statistics are computed on the cumulative histogram and only the light portion is examined. We can also assume the background is not a gradient with increasing darkness (this rare case would require a manual override button on the machine).

Thus, for simplification we assume that the statistical properties of the variation in the background's lightness are nearly constant along a scan. Consequently we can view the variation in the background's lightness as a stationary stochastic process. *Stationary* means that the probabilistic structure of different portions of the scanned document is the same.[3] Thus, the probability distribution of the background's stimulus luminance is the same at all the scan lines and the joint distribution of the background's stimulus luminances of pairs of scan lines distant h apart is always the same, etc. For our simple purpose of just determining at which scan lines to re-evaluate the exposure (the paper's tone level) this is an adequate approximation.

Because the process is stationary, the arrival times are exponentially distributed. Drawing on the previously mentioned experience (feeding paper skewness and paper insertion loss) putting the exponential distribution's parameter at 2 for the first 40 scan lines seems to be adequate. We tried a number of values for the parameter of the subsequent scan lines and obtained good results with the value 50. For the case of facsimile applications, ITU-T Recommendation T.4[8] specifies a maximum paper insertion loss of 4 mm and a maximum vertical loss caused by skew of ± 1.8 mm. At a resolution of 15.4 line/mm (400 lpi) this corresponds to 62 scan lines, so our numbers are somewhat optimistic (all our experiments were at 200 lpi and 300 lpi).

ITU-T Recommendation T.4 only specifies a minimum scan length (297 mm) but no maximum. In the case of a long document, it is safe to reuse the same random numbers to generate new scan line numbers for the exposure evaluation. A different pseudo-random number generator for exponentially distributed numbers could have been used and a different seed number could have been used. Each would produce a different but statistically equivalent set. It is also clear that a larger number of values can easily be changed by modifying the loop conditions in the program generating the scan line numbers.

It is important to notice that the generated table is fixed and each page will reset the index to the first element. This will ensure that when a page is re-scanned, the result will be the same if the page location in the mechanism is the same. For the statistical theory to be correct in this application, we only need to generate one table. We would only have to generate a new table for each scan, if we would like to compute some statistic across pages, which we do not. We only want to estimate the correct exposure for each individual page.

5.4 Tone reproduction

5.4.1 Fundamentals

As stated at the beginning, the aim of the dynamic exposure control mechanism is just to align the gamut of the original image with the dynamic range of the scanner system. An image's appearance depends on the shape of the *tone reproduction curve* (TRC). *Tone reproduction* is the relationship between the brightnesses in the original scene and those in the picture. Our method allows the designer to set the parameters for a good TRC.

The ability to design a tone reproduction curve is unique to digital image processing. In analog systems the TRC is a curve called the *characteristic curve* and is fixed for a given process. The manipulation possibilities are limited by the physical and chemical processes involved.

Many design methods for tone reproduction curves have been described in the literature since Leonardo da Vinci's chiaroscuro. In the following we will briefly discuss some practical considerations. Bartleson, Breneman and Hunt conducted extensive studies on the apparent contrast in black and white and color reproductions. The main paper [2] concludes that the relative brightnesses of a reproduction should match those of the original; regardless of the viewing conditions, the perceived brightness scale should be proportional to that which would correspond to the original.

Because tone reproduction refers to the appearance of images, tone reproduction curves are discussed in a perceptually linear space rather than in a space that is linear in intensity. Examples of spaces linear in intensity are the red, green, blue channels in scanners and display monitors. Examples of approximately perceptually linear spaces are optical density and $L*$.

Much work has then gone into scaling brightness and modeling the viewing conditions. An important result is that an image—which is related to complex field data—acts as its own induction field [1]; hence, the ordinary simple-field data of basic colorimetry cannot be used to compute gradients. An extensive overview of the subject can be found in Chapter 5 "Tone Reproduction and Color Balance" in Yule [10] and a more recent account in Hunt's book. [7]

Many scanners generate NTSC/RGB, but as we saw earlier this signal may or may not be gamma-corrected. If it is gamma-corrected the signal is approximately perceptual, otherwise the signal does not correlate well with perceived lightness.

5.5 TRCs in practice

The TRC that is stored into the scanner RAM (see Fig. 1) or in the host for the tone correction must be appropriate for the data type used. An easy mistake is to apply on intensity quantities curves designed for a perceptual quantity. In that case the TRC does therefore not adjust the tone correction, because no statement can be made about the perceived tonal rendition. It is more correct to state that the TRC step remaps the tones in the image to better take advantage of a device’s dynamic range; this impacts the tone reproduction, but we cannot quantify it for a human observer.

For reference, Fig. 4 shows the identity tone map, where the tone levels are left unchanged. In principle, this map works in the worst case situation, where the tone gamut and the dynamic range coincide and it acts as a pass-through. As Fig. 3 suggests, this is rarely the case and the tone gamut from the CCD is smaller than the display range. Few systems will use this TRC in practice.

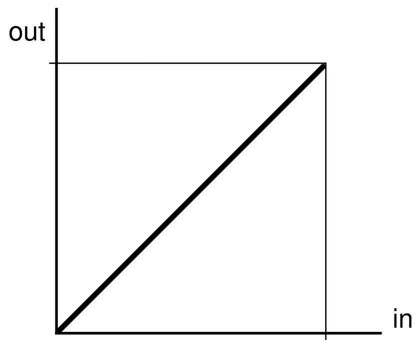

Figure 4. Identity tone map.

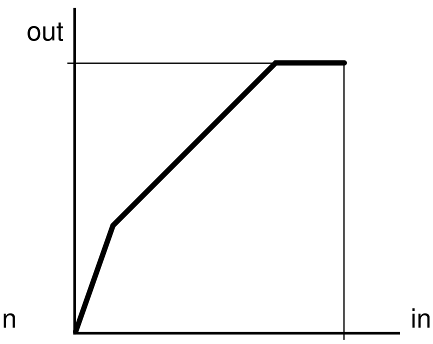

Figure 5. Offset tone map.

5.5.1 Tone correction by offset manipulation

Fig. 5 shows a tone correction by offset: A fixed value is added to the tone levels so that all levels higher than the exposure level (the level of the background) are mapped into the maximum value. The slope or *gradient* of the curve is unchanged, signifying that the *macro contrast* or *apparent contrast* of the image is not changed with respect to what is delivered by the sensors. To keep the black portion of the image (*e.g.*, the text in a document) black, a knee has to be introduced at a threshold value so that the darkest portion of the image remains dark. The disadvantage of a curve like Fig. 5 is that the dark portions of the image have a lower contrast, so that they are overemphasized with respect of the remainder of the image, yielding the appearance, in a printed hardcopy, of a poor ink coverage.

5.5.2 Tone correction by gain manipulation

This problem can be solved by using a gain instead of an offset to “push” the background level outside the dynamic range, as this is shown in Fig. 6. This method stretches the tone gamut to fit the dynamic range and is often used in practice.

At first this method might seem objectionable, because as for example Eschbach[5] notes, the apparent contrast of the image is altered, in violation of the recommendations by Bartleson et al. In practice this method yields acceptable results because the increased contrast compensates for the contrast reduction caused by veiling light[†] and flare [‡] in the scanner, printer, and/or display monitor. However, this reasoning would only be correct if the tone correction curve is applied in a perceptually linear space, which is not the case if it is applied to byte counts, especially when they are not gamma-corrected.

5.5.3 Tone correction by exponent manipulation

As mentioned in Section 1.1.1, "Fechner's law repealed", one correction to Stevens' law is to apply an exponential term. Empirically, we can state that if the tone gamut does not fit well in the dynamic range, the image looks under-exposed or over-exposed and we can change the emphasis of the dark or light tones.

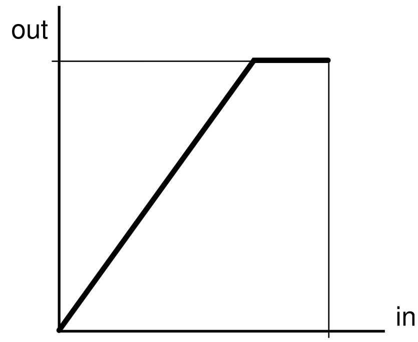

Figure 6. Gain tone map

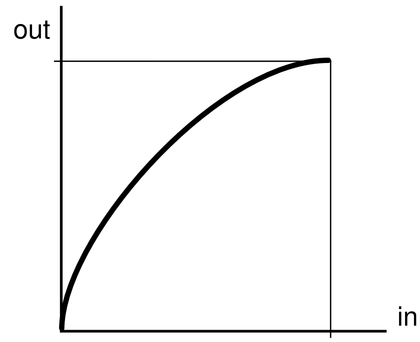

Figure 7. Exponential tone map

Fig. 7 illustrates Eschbach's method [5] in which an exponential correction is applied. Although the image's appearance is improved over the use of the identity TRC in Fig. 4, this TRC has the objectionable property that the linear proportion of the tone values in the image are destroyed, thus altering the expression of the image. This is particularly the case when the operation is performed in intensity space. When this TRC is applied to intensity data, we can state only that for an exponent $\gamma > 1$ that the tone levels in dark portion of the image are expanded while the light levels are compressed. If the exponent is changed to $\gamma < 1$, the effect is the opposite; the light areas are expanded while the dark areas are compressed.

5.5.4 The sigmoidal curve

In essence, the TRC is separated into three regions. They are called the *toe*, corresponding to dark values, the *shoulder*, corresponding to light values, and the *midtones* where most of the pictorial information in an image tends to reside. Fig. 8 shows a typical sigmoidal TRC. This shape is common in analog image processing and has the desirable property that a gradual degradation occurs at the extremes of the dynamic range, instead of the hard clipping typical for conventional digital systems such as gain manipulation (see Section 5.5.2).

The sigmoidal shape has the following properties:

†. Specular reflections that appear on the object viewed and that partially or wholly obscure the detail by reducing contrast.

‡. The effect of flare light is to reduce contrast appreciably in dark areas, but only slightly in light areas.

1. The toe is designed to cover the shadow portion in the image. The low slope keeps dark colors dark. For example, this ensures that black text in documents stays black and can easily be processed—for example—by an optical character recognition (OCR) stage in the receiving system because the edges remain sharper.
2. The mid-tone region is kept straight, so that the reproduction of these important tone values is linear when the quantities correlate to relative brightness. The slope of this curve is called *gamma* or *tone reproduction gradient*. We prefer to use the latter term to avoid confusion with the gamma used to designate the exponent in an exponential TRC. The gradient value is chosen so that it compensates for the flare in the system and for the apparent contrast reduction induced by the viewing conditions (*e.g.*, light, dim, or dark surround).
3. The shoulder tapers off so that the variability in the background's tone level can be controlled while some detail can be preserved in the highlights of the image.

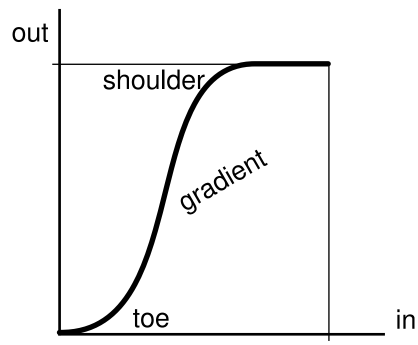

Figure 8. Sigmoidal tone map.

Edgar[4] has developed an analytical method for determining the parameters in a particular sigmoidal tone map. After the entire document has been scanned, the image's histogram is analyzed. The histogram is normalized to establish white-black references. Next, peaks attributable to image artifacts are limited. A second order best fit regression of the histogram is then performed and used to correct the image.

Continuing with Edgar, in the quadratic form $A + Bx + Cx^2$ the first term A is related to the number of pixels in the image and is not interesting. The coefficient B of the linear term is correlated with contrast (or tone reproduction gradient). For a low key image (high gradient and lots of dark pixel values) $B < 0$, while for a high key image (low gradient and lots of light pixel values) $B > 0$. The coefficient C of the quadratic term is correlated with saturation, which is characterized by the toe and shoulder of the TRC. Depending on the sign of coefficient B, $C > 0$ indicates that the shadows, the highlights, or both are compressed by saturation.

5.6 Application to exposure control

The dynamic exposure control method described in this paper decouples the design of the TRC from the actual data representing an image. The statistical analysis on the data is used to determine the parameters for a TRC designed for optimal system performance. Examples of the parameters are the position and curvature of the toe and shoulder of the TRC.

To ensure that the parameters change only slowly during the scan, each time the parameters are changed the new refined values based on the larger portion of the image are averaged with the previous parameter values. This increases the robustness of the method.

6 Conclusion

We have described a method to fully utilize the data-path bandwidth in a sheet-fed scanner. Our method has the advantage that instead of examining the signals pertinent to an image, statistical methods are used to estimate the tone level of the paper or other background that constitutes the substrate carrying an image or document. In other words, instead of an unqualified examination of CCD signals, we extract a semantic information primitive from the data.

Knowledge of the semantics decouples the tone reproduction characteristic from the exposure control necessary to map the images tone range into the scanner's dynamic range designed for a worst case situation. Our method is incremental: the statistical data is gathered during the scan. While the scan progresses, the estimate is refined based on the increased amount of data available from the accumulated histogram.

References

1. C.J. Bartleson, “Optimum Image Tone Reproduction,” *Journal of the SMPTE*, **84**, 613–618, 1975.
2. C.J. Bartleson and E.J. Breneman, “Brightness Perception in Complex Fields,” *Journal of the Optical Society of America*, **57**, 7, 953–957, 1967.
3. D.R. Cox and H.D. Miller, “*The Theory of Stochastic Processes*,” Chapman and Hall, London (United Kingdom), 1965.
4. A.D. Edgar, *System and method for automatic image saturation, gamma, and exposure correction in a digitizing video capture system*, United States Patent 5,265,200, 1993.
5. R. Eschbach, *Image-dependent exposure enhancement,* United States Patent 5,414,538, 1995.
6. I. Furukawa et al., “Super High Definition Image Digitizing System,” *ICASSP '92: Acoustics, Speech & Signal Processing Conference*, **III**, 529–532, 1992.
7. R.W.G. Hunt, “*The Reproduction of Colour in Photography, Printing & Television*,” 4th edition, Fountain Press, Tolworth (United Kingdom), 1987.
8. International Telecommunication Union, Telecommunication Standardization Sector, Terminal Equipments and Protocols for Telematic Services, Recommendation T.4: *Standardization of Group 3 Facsimile Apparatus for Document Transmission*, Geneva (Switzerland), 1993.
9. M. Ito, *White reference data generating unit applied to shading correction system*, United States Patent 5,317,421, 1994.
10. J.A.C. Yule, “*Principles of Color Reproduction*,” Wiley & Sons, New York, 1967.